

RANDOM PERTURBATION TO THE GEODESIC EQUATION

XUE-MEI LI

ABSTRACT. We study random perturbations to the geodesic equation. If the velocity of the geodesic with unit initial velocity is stirred sufficiently uniformly, the solutions, after suitable rescaling, converge to a Brownian motion scaled by $\frac{8}{n(n-1)}$ where n is the dimension of the state space.

1. INTRODUCTION

Let M be a complete smooth Riemannian manifold, TM its tangent bundle and T^*M its cotangent bundle. A geodesic $(x(t), t \in [0, 1])$ is a solution to the geodesic equation $\frac{d}{dt}\dot{x}^k(t) = -\sum_{i,j} \Gamma_{ij}^k(x(t))\dot{x}^i(t)\dot{x}^j(t)$, where the functions $\{\Gamma_{ij}^k\}$ are the Christoffel symbols. A geodesic is intuitively the motion of a free particle that minimises the energy function $E(x) = \frac{1}{2} \int_0^1 |\dot{x}(t)|^2 dt$; they are critical points of E . The velocity of a geodesic lives in the tangent bundle, but if we identify the tangent bundle with the cotangent bundle, the geodesic flow is the Hamiltonian flow on the cotangent bundle for the Hamiltonian function $H(x, y) = \frac{1}{2}|y|_x^2$, $(x, y) \in T^*M$. Let (O, x) be a local coordinate chart for M , where O is an open set of M and by abuse of notation $x : O \rightarrow \mathbb{R}^n$ is a diffeomorphism to its image and $x = (x^1, \dots, x^n)$. Then (x, y) is the induced coordinate map for T^*M , and (x, y) represents the cotangent vector $\sum_i y_i dx_i$. Let (g^{ij}) denote the inverse matrix to the Riemannian metric (g_{ij}) then $H(x, y) = \frac{1}{2} \sum_{i,j} g^{ij}(x) y_i y_j$. Let $\omega = \sum_i dx_i \wedge dy_i$ be the non-degenerate 2-form and define a vector field X on T^*M by $\iota_X \omega = dH$ where ι denotes interior product. The solution flow to $\dot{x}(t) = X(x(t))$ are geodesics, to see this more clearly

keywords. random perturbation, stochastic differential equations, homogenisation, geodesics.

Mathematics Subject Classification (2000). 60H10, 58J65, 37Hxx, 53B05

This research was supported by the EPSRC grant EP/E058124/1.

we will write the equation in local coordinates. First

$$dH = \frac{1}{2} \sum_{i,j,k} \frac{\partial g^{i,j}}{\partial x_k} y_i y_j dx_k + \sum_{i,j} g^{ij} y_i dy_j.$$

If $X = \sum_k f_k \frac{\partial}{\partial x_k} + \sum_k h_k \frac{\partial}{\partial y_k}$, then $\iota_X \omega = \sum_k f_k dy_k - \sum_k h_k dx_k$. This means that X has the expression

$$X = \sum_k \sum_i g^{ik} y_i \frac{\partial}{\partial x_k} - \sum_k \frac{1}{2} \sum_{i,j} \frac{\partial g^{i,j}}{\partial x_k} y_i y_j \frac{\partial}{\partial y_k}.$$

Let (x_t, y_t) denote the integral curve of X , then

$$\dot{x}^k = \sum_j g^{kj} y_j, \quad \dot{y}_k = -\frac{1}{2} \sum_{i,j} \frac{\partial g^{ij}}{\partial x_k} y_i y_j.$$

Next we differentiate \dot{x}^k once more, transform y_k to \dot{x}^k 's by raising the indices, apply the formula for Christoff symbols in terms of $(g_{i,j})$, and we see that this is indeed the geodesic equation.

Our formulation of the perturbation to the geodesics is best described by perturbation to an ODE on the frame bundle, see §2C for the passage from one to the other. A geodesic is the projection of a horizontal flow from the bundle of orthonormal frames of M to M . Let u stand for an orthonormal frame at a point x of M , i.e. an orthonormal basis of $T_x M$. Let us denote by π the map that takes a frame u of $T_x M$ to the point $x \in M$. For $e_0 \in \mathbb{R}^n$, let $H_u(e_0)$ be the horizontal lift of $u(e_0)$. Let $(u_t^{e_0})$ be the solution to

$$\dot{u}(t) = H_{u_t}(e_0), \quad u(0) = u_0,$$

then $\pi(u_t^{e_0})$ is the geodesic with initial velocity $u_0(e_0)$ and initial point $\pi(u_0)$. We perturb this ODE in directions that are transversal,

$$\dot{u}_t^\epsilon = H_{u_t^\epsilon}(e_0) + G^\epsilon(u_t^\epsilon),$$

where G^ϵ is to be specified, and conclude that there is an effective motion which projects to a scaled Brownian motion on M with a factor $\frac{8}{n(n-1)}$. The perturbation G^ϵ involves a Gaussian noise, so techniques involving parallel transport along a stochastic process and horizontal lifts of non-smooth curves will be used. Stochastic parallel translations goes back to Itô in 1962, [24, 26] where piecewise approximation was used for the construction, followed by Dynkin [8]. The canonical construction using a Stratonovich SDE on the orthonormal frame bundle can be found in Eells-Elworthy [9]. See also Malliavin [34]. Horizontal stochastic processes have been used in connection with the following topics: horizontal lifts of semi-martingales, construction of canonical

Brownian motions, Malliavin calculus, construction of line integrals, the geometry of diffusion operators. This study is an extension to and an application of the last mentioned topic.

Let $N = \frac{n(n-1)}{2}$. If $A \in \mathfrak{so}(n)$ we denote by A^* the fundamental vertical vector field on OM determined by right actions of the exponentials of tA , see (2.1) below. We denote by Δ^G the Laplacian on G , and Δ_H the horizontal Laplacian on OM . If X is a vector field, we denote by L_X Lie differentiation in the direction of X . Let $A_k \in \mathfrak{so}(n)$ and $\mathcal{L}_G = \frac{1}{2} \sum_{k=1}^N L_{gA_k} L_{gA_k}$, at $g \in G$. We denote by \circ Stratonovich integration. Let us fix a time $T > 0$ and the stochastic processes are on the interval $[0, T]$. Let e_0 be a unit vector in \mathbb{R}^n . Let ρ be the Riemannian distance function on M and ∇ the Levi-Civita connection.

Theorem 1.1. *Let M be a complete Riemannian manifold of dimension $n > 1$ of positive injectivity radius. Suppose that there are positive numbers C, a such that $\sup_{\rho(x,y) \leq a} |\nabla d\rho|(x, y) \leq C$. Let $x_0 \in M$ and $u_0 \in \pi^{-1}(x_0)$. Let $\{A_1, \dots, A_N\}$ be an o.n.b of \mathfrak{g} , and $\bar{A} \in \mathfrak{g}$. Let (u_t^ϵ) be the solution to the SDE*

$$(1.1) \quad \begin{cases} du_t^\epsilon = H_{u_t^\epsilon}(e_0)dt + \frac{1}{\sqrt{\epsilon}} \sum_{k=1}^N A_k^*(u_t^\epsilon) \circ dw_t^k + \bar{A}^*(u_t^\epsilon)dt, \\ u_0^\epsilon = u_0. \end{cases}$$

The SDE is conservative. Let $x_t^\epsilon = \pi(u_t^\epsilon)$ and let (\tilde{x}_t^ϵ) be the horizontal lift of (x_t^ϵ) to OM through u_0 . Then

- (1) *The processes $(x_{\frac{t}{\epsilon}}^\epsilon)$ and $(\tilde{x}_{\frac{t}{\epsilon}}^\epsilon)$ converge in law, as $\epsilon \rightarrow 0$.*
- (2) *The limiting law of $(x_{\frac{t}{\epsilon}}^\epsilon)$ is independent of e_0 . It is a scaled Brownian motion with generator $\frac{4}{n(n-1)}\Delta$. The limiting probability distribution of $(\tilde{x}_{\frac{t}{\epsilon}}^\epsilon)$ is that associated to the generator $\frac{4}{n(n-1)}\Delta_H$.*

If ' $\epsilon = \infty$ ', the equation is the 'geodesic equation': $\dot{x}_t^\infty = u_t^\infty e_0$. The perturbation to the geodesic is exerted only through the perturbation of its velocity. Since u_t is a linear isometry, the velocity of the motion is always unitary. The effective motion is due to the fast rotation in the velocity field. The perturbed geodesic has rapid changing directions and we expect to see a jittering motion and indeed we obtain a scaled Brownian motion in the limit if the rotational motion is elliptic.

This agrees with the philosophy in Bismut [4], that $\ddot{x} = \frac{1}{T}(-\dot{x} + \dot{w})$ interpolates between classical Brownian motion ($T \rightarrow 0$) and the geodesic flow ($T \rightarrow \infty$). We also note limit theorems on line integrals of

the form $\int_0^t \phi(dx_s)$, where ϕ is a differential form and (x_s) is a suitable process such as a Brownian motion. See Ikeda [22] and Ikeda-Ochi [23]. A central limit theorem is proven to be valid for line integrals along geodesic flows by Manabe-Ochi [35], where they used symbolic representations of geodesic flows. A related work is that by Pinsky [39], where a piecewise geodesic with a Poisson-type switching mechanism is shown to converge to the horizontal Brownian motion.

The scaling limit is consistent with the central limit theorem for geodesic flows $\theta_s(v) = (\gamma_t(x, v), \dot{\gamma}_t(x, v))$ on the unit tangent bundle, where $(\gamma_t(x, v))$ denotes the geodesic with initial value $(x, v) \in STM$. Let us assume that M is a manifold of constant negative curvature and of finite volume. If f is a bounded measurable function on the unit tangent bundle, centred with respect to the normalised Louville measure \mathbf{m} , then the central limit theorem states that there is a number σ with the property that

$$\lim_{t \rightarrow \infty} \mathbf{m}\left\{\xi : \frac{\int_0^t f(\theta_s(\xi)) ds}{\sigma \sqrt{t}} \leq a\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a e^{-\frac{y^2}{2}} dy.$$

See Sinai [41], Ratner [40], and Guivarch-LeJan [18], and Enriquez-Franchi-LeJan [15] for further developments. These results explore the chaotic nature of the deterministic dynamical system on manifolds of negative curvatures.

In the homogenisation literature, the following work are particularly relevant: Khasminskii [28, 20], Nelson [36], Borodin, Freidlin [5], Freidlin, Wentzell [17], and Bensoussan, Lions, Papanicolaou [2]. In terms of scaling limits in manifolds we refer to Li [32] for averaging of integrable systems and to Gargate, Ruffino [27] for averaging on foliated manifolds. See also [33] for an earlier work on the orthonormal frame bundle. We also refer to Dowell [7] for a scaling limit of Ornstein-Uhlenbeck type and to Bismut [4] on Hypoelliptic Laplacians and orbital integrals.

Open Questions. (1) We have a rate estimate in Lemma 3.4, but we do not know the rate of convergence in Theorem 1.1. Can an estimate be obtained? (2) The local uniform bound on $\nabla d\rho$ is only used in Lemma 3.2 for the proof of tightness. This bound can be weakened, for example replaced by a local uniform control over the rate of growth of the norms of $\frac{\nabla d\rho}{\rho}$ and $\frac{\nabla \rho}{\rho}$. For the purpose of this article it is too long to be included, however it is worth further study in view of stochastic completeness of a Riemannian manifold. A manifold is stochastically complete if the Brownian motion is complete, i.e. has infinite life time. A geodesically complete manifold is not necessarily

stochastically complete; this was pointed out in Azencott [1] where the author studied negatively curved manifolds and noted that if the sectional curvature decays at infinity faster than ρ^{2+} , the manifold is stochastically incomplete. There have since been many results on the stochastic completeness. They are mostly in terms of the global decay of the Ricci curvature at infinity and the volume growth of a ball of radius r . The Brownian motion constructed in Theorem 1.1 will be automatically complete. The conditions in Theorem 1.1 appear to be related to the uniform cover criterion on stochastic completeness and could be studied in connection with that in Li [31]. If the manifold has a positive injectivity radius a , every point in the manifold is contained in a chart (O, x) with O contains a ball of radius a . A ‘uniform cover’ type condition for an SDE $dx_t = \sum_k \sigma_k(x_t) \circ dw_t^k + \sigma_0(x_t)dt$ is on the size of $|\phi_*(\sigma_k)|$ or on how fast does a Brownian motion escapes the ball. Such bounded local coordinate method goes back to Itô [25] and was fully developed in Elworthy [12], and see also Clark [6]. This method evolved into ‘weak uniform covers’ in Li [31] where it was shown to be an effective criterion for the non-explosion and for the C_0 -property of the semi-groups. Also much of the work in this article is valid for a connection ∇ with torsion; the horizontal tangent bundle and Δ_H will then be induced by this connection with torsion. The effect of the torsion will generally lead to an additional drift to the Brownian motion downstairs. In this case the geodesic completeness of the manifold M may no longer be equivalent to the metric completeness of the metric space (M, ρ) .

2. PRELIMINARIES

A. A frame u is an ordered basis of $T_x M$. We denote by FM the set of all frames on M and π the map that takes the frame u to the point $x \in M$. Let $\pi^{-1}(x) = \{u \in FM : \pi(u) = x\}$. We call FM the bundle of frames of M . Let $u = \{u_1, \dots, u_n\}$ be a frame, where $u_i \in T_x M$. If (O, x) is a coordinate system on M , $u_i = \sum_j u_i^j \frac{\partial}{\partial x_j} \Big|_x$. This gives a coordinate map on FM . The map (x, u_i^j) is a homeomorphism from $\pi^{-1}(O)$ to $(x(O), GL(n, \mathbb{R}))$. We may consider the subspace of FM that consists of bases of $T_x M$ that are orthonormal, w.r.t. the given Riemannian metric, in which case we have the bundle of orthonormal frames. We denote the orthonormal frame bundle by OM . The orthonormal frame bundle is a fibre bundle with group $O(n)$. If the manifold is oriented we may consider a connected component SOM .

If we identify a frame u by the transformation $u : \mathbb{R}^n \rightarrow T_x M$, FM , OM and SOM are principle bundles with fibres $GL(n, \mathbb{R})$, $O(n)$, and

$SO(n)$ respectively. For ease of notation we denote by P one of the bundles and G one of the groups and \mathfrak{g} its Lie algebras. The group G acts on the right: if u is a frame and $g \in G$ then ug is another frame. For $g \in G$ let R_g denotes the right action on G or on P . We use the following conventions for G the orthogonal group or the special orthogonal group. If $A, B \in \mathfrak{so}(n)$, we define $\langle A, B \rangle = \text{tr } AB^T$ which is bi-invariant and we define the Riemannian metric at $T_g G = \{A : gA^T + Ag^T = 0\}$ to be $\langle gA, gB \rangle_g = \langle A, B \rangle$.

A tangent vector v in P is vertical if $T\pi(v) = 0$. We introduce a family of vertical vector fields. If A belongs to the Lie algebra \mathfrak{g} , we denote by $\exp(tA)$ the exponential map. If u is a frame, the composition $u \exp(tA)$ is again a frame in the same fibre. We define the fundamental vertical vector fields associated to A by A^* ,

$$(2.1) \quad A^*(u) = \left. \frac{d}{dt} \right|_{t=0} u \exp(tA).$$

B. Suppose that we are given an Ehresmann connection : $T_u P = HT_u P \oplus VT_u P$, so every tangent vector on FM has a horizontal component and a vertical component. The horizontal space is right invariant: $(R_a)_* HTP = HTP$ and the projection π induces an isomorphism between HTP and TM .

An Ehresmann connection determines and is determined by parallel translation. A piecewise C^1 curve γ on P is horizontal if the one sided derivatives $\dot{\gamma}(\pm)$ are horizontal for all t . If c is a C^1 curve on M there is a horizontal curve \tilde{c} on P such that \tilde{c} covers c , i.e. $\pi(\tilde{c}(t)) = c(t)$. In fact $\tilde{c}(t)$ is the family of orthonormal frames along c that are obtained by parallel transporting the frame $\tilde{c}(0)$. We say that \tilde{c} is a horizontal lift of c . We will assume that the parallel translation is induced by the Levi-Civita connection. If \tilde{c} is a horizontal lift of c , the translation of \tilde{c} by $g \in G$, $R_g \circ \tilde{c}^*$, is also a horizontal curve. If we fix $\tilde{c}(0) = u_0 \in P$, there is only one horizontal lift with $\tilde{c}(0) = u_0$. Let $v = \dot{\tilde{c}}(0)$ and $u = \tilde{c}(0)$. We define the horizontal lift of $c(0)$, $\mathfrak{h}_u(v)$, to be $\dot{\tilde{c}}(0)$. If c is a solution of a stochastic differential equation we use the concept and theory of horizontal lifts in Eells-Elworthy [9, 11], Malliavin [34], Elworthy [10] and Emery [14]. We follow the notation in [10].

To each $e \in \mathbb{R}^n$ we associate a special horizontal vector field, the basic vector field $H_u(e) = \mathfrak{h}_u(ue)$. They satisfy $\pi_*(H_u(e)) = ue$. We will introduce a metric on OM such that π is an isometry between $H_u TOM$ and $T_{\pi(u)} M$, and on each fibre it is the bi-invariant metric on the Lie algebra. If $\{e_1, \dots, e_n\}$ is an orthonormal basis of \mathbb{R}^n then $\{H_u(e_1), \dots, H_u(e_n)\}$ is an orthonormal basis for the horizontal tangent subspace HTP at u .

C. We describe the relation between horizontal equations on frame bundles and the geodesic flows. The tangent bundle TM is the fibre bundle associated with the principal fibre bundle P with fibre \mathbb{R}^n . The total space is $P \times \mathbb{R}^n / \sim$ where the equivalent class is determined by $[u, e] \sim [ug^{-1}, ge]$, any $g \in G$. Fix a unit vector $e_0 \in \mathbb{R}^n$. Let H be the isotropy group at e_0 of the action of G on \mathbb{R}^n . Each element $v \in TM$ has a representation $[u, e_0]$ in $P \times \mathbb{R}^n$ and it is unique up to right translation by elements of H . We may identify $P \times \mathbb{R}^n / \sim$ with the quotient bundle P/H , whose element containing u is the equivalence class of elements of the form $ug, g \in H$. Let α be the associated map:

$$\alpha_{e_0} : u \in P \rightarrow ue_0 \equiv [u, e_0] \in TM.$$

The differential $D\alpha_{e_0}$ induces a map from $T_u P$ to $T_{ue_0} TM$. Any vector field W that is invariant by right translations of elements of H induces a vector field on TM . If $v = ue'_0 = ue_0$ there is $g \in G$ with $e'_0 = g^{-1}e_0$. Set $u' = ug$. Since $\alpha_{e_0}(u) = \alpha_{e'_0}(R_g u)$,

$$D_u \alpha_{e_0}(W(u)) = D_{u'} \alpha_{e'_0} D R_g(W(u)) = D_{u'} \alpha_{e'_0}(W(u')),$$

and the map $W \in \Gamma TP \mapsto D_u \alpha_{e_0}(W(u)) \in \Gamma TTM$, is independent of the choices of e_0 . If $W(u) = H_u(e_0)$, the induced vector field X on the tangent bundle TM is a geodesic spray, i.e. in local co-ordinates $X(x, v) = (x, v, v, Z(x, v))$ and $Z(x, sv) = s^2 Z(x, v)$. This corresponds to the geodesic equation on TM : $dv_t^k = -\Gamma_{ij}^k(\sigma_t) v_t^i v_t^j$, $\dot{\sigma}_t = v_t$, $\sigma(0) = \pi(u_0)$, $v(0) = u_0 e_0$. A vector field on P that is horizontal and invariant under translation by the action of G projects to a vector field on the base manifold. It is worth remarking that $H(e_0)$ does not project to a vector field on M .

D. A basic object we use in our computation is the connection 1-form ϖ on FM corresponding to the Ehresmann connection. We follow closely the notation in Kobayashi-Nomizu [30]. A connection form assigns a skew symmetric matrix to every tangent vector on FM and it satisfies the following conditions: (1) $\varpi(A^*) = A$ for all $A \in \mathfrak{g}$; (2) for all $a \in G$ and $w \in FM$, $\varpi(R_{a*} w) = Ad(a^{-1})\varpi(w)$. We recall that $R_{a*}(A^*) = (Ad(a^{-1})A)^*$ for all $a \in G$. It is convenient to consider horizontal tangent vectors on P are elements of the kernel of ϖ . If $\{A_1, \dots, A_N\}$ is a basis of \mathfrak{g} , then the horizontal component of a vector w is $w^h = w - \sum_j \varpi(w) A_j^*$.

The connection 1-form ϖ is basically the set of Christoffel symbols. Let $E = (E_1, \dots, E_n)$ be a local frame; we define the Christoffel symbols relative to E by $\nabla E_j = \sum_{ki} \Gamma_{ij}^k dx_i \otimes E_k$. Let θ^i be the set of dual differential 1-forms on M to $\{E_i\}$: $\theta^i(E_j) = \delta_{ij}$. We define $\omega_i^j = \Gamma_{kj}^i \theta^k$.

Then $d\theta^i = -\sum_k \theta^k \wedge \omega_k^i$ and ω_i^j 's measure the change of the dual forms θ^i in the direction of E_j . Let A_i^j be a basis of \mathfrak{g} , to each moving frame we associate a 1-form, $\omega = \sum_{i,j} \omega_j^i A_i^j$, on M . If (O, x) is a chart of M and $s : O \rightarrow OM$ is a local section of OM , let us denote by ω_s the differential 1-form given above, then $\varpi(s_*v) = \omega_s(v)$. Conditions (1) and (2) are equivalent to the following: if $a : U \rightarrow G$ is a smooth function,

$$\varpi((s \cdot a)_*v) = a^{-1}(x)da(v) + a^{-1}(x)\varpi(s_*v)a(x).$$

This corresponds to the differentiation of $s \cdot a$.

E. Let us work in a coordinate chart (O, x) . Let $c(t)$ be a curve and $\tilde{c}(t)$ a horizontal lift of $c(t)$ whose column vectors $\{\tilde{c}_l\}$ is a frame. Then $\tilde{c}(t)$ satisfies:

$$\frac{\partial \tilde{c}_l^k}{\partial t} + \sum_{i=1}^n \frac{\partial c^i}{\partial t} \Gamma_{ij}^k \tilde{c}_l^j = 0.$$

Take $c(t) = (0, \dots, t, \dots, 0)$, where the non-zero entry is in the i -place. We obtain the principal part of the horizontal lift of $\frac{\partial}{\partial x_i}$:

$$\left(\mathfrak{h}_{\tilde{c}(0)} \left(\frac{\partial}{\partial x_i} \right) \right)_l = \left(\frac{\partial \tilde{c}}{\partial t} (0) \right)_l = - \left(\sum_j \Gamma_{ij}^1 u_l^j, \dots, \sum_j \Gamma_{ij}^n u_l^j \right)^T.$$

The horizontal space at u is spanned has a basis

$$\left\{ \left(\frac{\partial}{\partial x_i}, - \left(\sum_{j=1}^n u_a^j \Gamma_{ij}^b \frac{\partial}{\partial u_a^b} \right) \right) \right\}.$$

Example 2.1. Let $H^2 = \{(x_1, x_2) : x_2 > 0\}$ be the the hyperbolic plane. It has a global chart and $g_{ij} = \frac{1}{(x_2)^2} \delta_{ij}$. Its non-zero Chrsitoffel symbols are:

$$\Gamma_{12}^1 = \Gamma_{21}^1 = -\frac{1}{y}, \quad \Gamma_{22}^2 = -\frac{1}{y}, \quad \Gamma_{11}^2 = \frac{1}{y}.$$

The total space of the orthonormal frames is a product space. We let u denote the principal part of an frame. For $u = (u_1, u_2)$ with $u_i = (u_i^1, u_i^2)^T$ and $\pi(u) = x = (x_1, x_2) \in H^2$, let us define

$$X_1(u) = \frac{1}{x_2} \begin{pmatrix} -u_1^2 & -u_2^2 \\ u_1^1 & u_2^1 \end{pmatrix}, \quad X_2(u) = -\frac{1}{x_2} \begin{pmatrix} u_1^1 & u_2^1 \\ u_1^2 & u_2^2 \end{pmatrix}, \quad A = \begin{pmatrix} 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 \end{pmatrix}.$$

Let us take $e_0 = e_1$ then $ue_1 = u_1$ and our SDE becomes:

$$\dot{x} = u_1$$

$$du = u_1^1 \cdot X_1(u)dt - u_1^2 \cdot X_2(u)dt + \frac{1}{\sqrt{\epsilon}} u A \circ dw_t.$$

In the equation we suppressed the t -variable as well as the superscript ϵ in the stochastic processes $(x^\epsilon(t), u^\epsilon(t))$. These are further subject to the following three constraints:

$$(u_1^1)^2 + (u_1^2)^2 = (x_2)^2, \quad (u_2^1)^2 + (u_2^2)^2 = (x_2)^2, \quad u_1^1 u_2^1 + u_1^2 u_2^2 = 0.$$

We hence have a system of 3 SDE's with a one dimensional driving Brownian motion (w_t) . The effective motion, obtained from the x process, is a scaled hyperbolic Brownian motion.

Example 2.2. Let us strip off the geometry and take a close look at the example of $M = \mathbb{R}^d$ with the trivial Riemannian metric. Then $FM = \mathbb{R}^n \times GL(n)$ and $OM = \mathbb{R}^n \times SO(n)$ are the trivial product bundles. The horizontal vectors in the tangent space of OM are those whose Lie-algebra component vanishes. We write a frame u as (x, g) . The horizontal lift at $u = (x, g)$ of a vector $v \in T_x \mathbb{R}^n$ is $((x, g), (v, 0)) \in (\mathbb{R}^n \times G) \times (T_x \mathbb{R}^n \times \mathfrak{g})$. To ease the notation we omit the trivial component of the horizontal lift, we have $H_u(e) = (ge, 0)$ and the equation $\dot{u}_t = H_{u_t}(e_0)$ is equivalent to $\dot{x}_t = g_t e_0, \dot{g}_t = 0, g_0 e_0 = v_0$. Let $A_k \in \mathfrak{so}(n)$, the perturbed system is

$$\dot{x}_t^\epsilon = g_t^\epsilon e_0, \quad \dot{g}_t^\epsilon = \frac{1}{\sqrt{\epsilon}} \sum_k g_t^\epsilon A_k \circ dw_t^k + g_t^\epsilon \bar{A} dt,$$

with $x_0^\epsilon = x_0$ and $g_0^\epsilon = I$, the identity matrix. We claim that $x_{\frac{t}{\epsilon}}^\epsilon$ converges. At first glance this equation appears to have the wrong scaling. Let us set $y_t^\epsilon = \frac{1}{\epsilon} x_{\epsilon t}^\epsilon$ and $\tilde{g}_t^\epsilon = g_{\epsilon t}^\epsilon$. The above equation is equivalent to

$$\dot{y}_t^\epsilon = \tilde{g}_t^\epsilon e_0, \quad \dot{\tilde{g}}_t^\epsilon = \sum_k \tilde{g}_t^\epsilon A_k \circ d\tilde{w}_t^k + \epsilon \tilde{g}_t^\epsilon \bar{A} dt$$

with $y_0^\epsilon = \frac{1}{\epsilon} x_0$ and $\tilde{g}_0^\epsilon = I$. Here $\{\tilde{w}_t^k\}$ is a family of independent Brownian motions and (g_t^ϵ) will be independent of ϵ . Let us assume that $\bar{A} = 0$, then $\{g_t\}$ is a reversible ergodic Markov process on G with the invariant measure the Haar measure. We can apply central limit theorems for additive functionals of Markov processes. Let $e \in \mathbb{R}^n$ we set $V^e(g) = \langle g e_0, e \rangle$ and set $Y^e(t) = \int_0^t V^e(\tilde{g}_s) ds$. It is easy to check that for each e , V^e satisfies the conditions in Theorem 1.8 and Corollary 1.9 in Kipnis-Varadhan [29], in particular $\int V^e dg = 0$. Hence

$\epsilon Y^\epsilon(\frac{t}{\epsilon^2})$ converges, and so does $x_t^\epsilon = \epsilon y_{\frac{t}{\epsilon^2}}^\epsilon$. See also Helland [21]. If the connection on $TO\mathbb{R}^n$ is not trivial, the non-zero Christoffel symbol will be involved. In this case the ergodic component and the non-ergodic component oscillate at speeds of different scale.

3. SOME LEMMAS

Lemma 3.1. *Let M be a geodesically complete Riemannian manifold. Let (u_t^ϵ) be the solution to the SDE (1.1). Let $x_t^\epsilon = \pi(u_t^\epsilon)$, which has a unique horizontal lift, \tilde{x}_t^ϵ , through $u_0 \equiv u_0^\epsilon$. Then*

$$\begin{aligned} \frac{d}{dt} \tilde{x}_t^\epsilon &= H_{\tilde{x}_t^\epsilon}(g_t^\epsilon e_0) \\ dg_t^\epsilon &= \frac{1}{\sqrt{\epsilon}} \sum_{k=1}^m g_t^\epsilon A_k \circ dw_t^k + g_t^\epsilon \bar{A} dt, \end{aligned}$$

where g_0^ϵ is the identity matrix. Consequently the SDE (1.1) is conservative.

Proof. By the defining properties of the basic horizontal vector fields, $\dot{x}_t^\epsilon = \pi_*(H_{u_t^\epsilon}(e_0)) = u_t^\epsilon e_0$. Since $u_t^\epsilon e_0$ has unit speed, the solution exists for all time if (u_t^ϵ) does, and

$$\frac{d}{dt} \tilde{x}_t^\epsilon = \mathfrak{h}_{\tilde{x}_t^\epsilon}(\dot{x}_t^\epsilon) = \mathfrak{h}_{\tilde{x}_t^\epsilon}(u_t^\epsilon e_0).$$

At each time t , the horizontal lift (\tilde{x}_t^ϵ) of the curve (x_t^ϵ) through u_0 and the original curve u_t^ϵ belong to the same fibre. Let g_t^ϵ be an element of G with the property that $u_t^\epsilon = \tilde{x}_t^\epsilon g_t^\epsilon$. Then g_0^ϵ is the identity matrix and

$$\frac{d}{dt} \tilde{x}_t^\epsilon = \mathfrak{h}_{\tilde{x}_t^\epsilon}(\tilde{x}_t^\epsilon g_t^\epsilon e_0) = H_{\tilde{x}_t^\epsilon}(g_t^\epsilon e_0).$$

If a_t is a C^1 path with values in G , $a_t^{-1} \dot{a}_t = \frac{d}{dr}|_{r=0} e^{ra_t^{-1} \dot{a}_t}$, its action on u gives rise to a fundamental vector field,

$$\left. \frac{d}{dt} \right|_t ua_t = \left. \frac{d}{dr} \right|_{r=0} ua_t a_t^{-1} a_{r+t} = (a_t^{-1} \dot{a}_t)^*(ua_t).$$

By Itô's formula applied to the product $\tilde{x}_t^\epsilon g_t^\epsilon$,

$$du_t^\epsilon = TR_{g_t^\epsilon} d\tilde{x}_t^\epsilon + (TL_{(g_t^\epsilon)^{-1}} \circ dg_t^\epsilon)^*(u_t^\epsilon).$$

Since right translation of horizontal vectors are horizontal, the connection 1-form vanishes on the first term and $\varpi(\circ du_t^\epsilon) = TL_{(g_t^\epsilon)^{-1}} \circ dg_t^\epsilon$.

We apply ϖ to the SDE for u_t^ϵ ,

$$\begin{aligned} dg_t^\epsilon &= TL_{g_t^\epsilon} \varpi(\circ du_t^\epsilon) = TL_{g_t^\epsilon} \varpi\left(\frac{1}{\sqrt{\epsilon}} \sum_{k=1}^N A_k^*(u_t^\epsilon) \circ dw_t^k + \bar{A}^*(u_t^\epsilon) dt\right) \\ &= \frac{1}{\sqrt{\epsilon}} \sum_{k=1}^m g_t^\epsilon A_k \circ dw_t^k + g_t^\epsilon \bar{A} dt. \end{aligned}$$

There is a global solution to the above equation. The ODE $\frac{d}{dt} \tilde{x}_t^\epsilon = H_{\tilde{x}_t^\epsilon}(g_t^\epsilon e_0)$ has bounded right hand side and has a global solution. It follows that $u_t^\epsilon = \tilde{x}_t^\epsilon g_t^\epsilon$ has a global solution. \square

The stochastic process (g_t^ϵ) is a Markov process on $SO(n)$ with infinitesimal generator

$$\mathcal{L}_G^\epsilon = \frac{1}{2\epsilon} \sum_{k=1}^N L_{gA_k} L_{gA_k} + L_{g\bar{A}}.$$

If $\{A_k\}$ is an orthonormal basis of \mathfrak{g} , then $\mathcal{L}_G^\epsilon = \frac{1}{2\epsilon} \Delta^G + L_{g\bar{A}}$. Let $\mathcal{L}_G = \frac{1}{2} \sum_{k=1}^N L_{gA_k} L_{gA_k}$ so $\mathcal{L}_G^\epsilon = \frac{1}{\epsilon} \mathcal{L}_G + L_{g\bar{A}}$.

Lemma 3.2. *Let M be a complete Riemannian manifold with positive injectivity radius. Suppose that there are numbers $C > 0$ and $a_2 > 0$ such that $\sup_{\rho(x,y) \leq a_2} |\nabla d\rho|(x,y) \leq C$. Let $T > 0$. The probability distributions of the family of stochastic processes $\{\tilde{x}_t^\epsilon, t \leq T\}$ is tight. There is a metric \tilde{d} on M such that $\{(\tilde{x}_t^\epsilon)\}$ is equi Hölder continuous with exponent $\alpha < \frac{1}{2}$.*

Proof. Let μ^ϵ be the probability laws of (\tilde{x}_t^ϵ) on the path space over OM with initial value u_0 , which we denote by $C([0, T]; OM)$. Since $\tilde{x}_0^\epsilon = u_0$ it suffices to estimate the modulus of continuity and show that for all positive numbers a, η , there exists $\delta > 0$ such that for all ϵ sufficiently small, see Billingsley [3] and Ethier-Kurtz [16],

$$P(\omega : \sup_{|s-t| < \delta} d(\tilde{x}_t^\epsilon, \tilde{x}_s^\epsilon) > a) < \delta\eta.$$

Here d denotes a distance function on OM . We will choose a suitable distance function. The Riemannian distance function $\tilde{\rho}(x, y)$ is not smooth in y if y is in the cut locus of x . To avoid any assumption on the cut locus of OM we construct a new distance function that preserves the topology of OM .

Let $2a$ be the minimum of 1, a_1 and the injectivity radius of M . Let $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ be a smooth concave function such that $\phi(r) = r$ when $r < a$ and $\phi(r) = 1$ when $r \geq 2a$. Let ρ and $\tilde{\rho}$ be respectively the

Riemannian distance on M and OM . Then $\phi \circ \rho$ and $\phi \circ \tilde{\rho}$ are distance functions on M and on \tilde{M} respectively. Then for $r < t$,

$$\phi^2 \circ \tilde{\rho}(\tilde{x}_\frac{t}{\epsilon}^\epsilon, \tilde{x}_\frac{r}{\epsilon}^\epsilon) = \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} D(\phi^2 \circ \tilde{\rho}(\tilde{x}_r^\epsilon, \cdot))_{\tilde{x}_s^\epsilon}(H_{\tilde{x}_s^\epsilon}(g_s^\epsilon e_0)) ds.$$

Since $H_{\tilde{x}_s^\epsilon}(g_s^\epsilon e_0)$ has unit length, from the equation above we do not observe directly a uniform bound in ϵ .

For further estimates we work with a C^2 function $F : OM \rightarrow \mathbb{R}$ to simplify the notation. Also the computations below and some of the identities will be used later in the proof of Theorem 1.1. Let $0 \leq r < t$,

$$(3.1) \quad F(\tilde{x}_\frac{t}{\epsilon}^\epsilon) = F(\tilde{x}_\frac{r}{\epsilon}^\epsilon) + \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} (DF)_{\tilde{x}_s^\epsilon}(H_{\tilde{x}_s^\epsilon}(g_s^\epsilon e_0)) ds.$$

Let $\{e_i\}$ be an orthonormal basis of \mathbb{R}^n . We define two sets of functions $f_i : OM \rightarrow \mathbb{R}$ and $h_i : G \rightarrow \mathbb{R}$:

$$f_i(u) = (DF)_u(H_u e_i), \quad \alpha_i(g) = \langle g e_0, e_i \rangle.$$

From the linearity of H_u we obtain the identity $H_u(g e_0) = \sum_{i=1}^n f_i(u) \alpha_i(u)$.

Since the Riemannian metric on $SO(n)$ is bi-invariant the Riemannian volume measure, which locally has the form $\sqrt{\det(g_{ij})} dx^1 \wedge \cdots \wedge dx^n$, is the Haar measure. Let dg be the Haar measure normalised to be a probability measure. Let \tilde{g} be a vector such that $\tilde{g} e_0 = -e_0$. Then $\int_G g(\tilde{g} e_0) dg = \int_G g(e_0) dg$. The integral of $g e_0$ with respect to the Haar measure on G vanishes. In particular $\int_G \alpha_i dg = 0$. On a compact Riemannian manifold the Poisson equation with a smooth function that is centred with respect to the Riemannian volume measure has a unique centred smooth solution. For each i , let $h_i : G \rightarrow \mathbb{R}$ be the smooth centred solution to the Poisson equation $\mathcal{L}_G = \alpha_i$. We apply Itô's formula to the function $f_i h_i$ and $r < t$,

$$\begin{aligned} f_i(\tilde{x}_\frac{t}{\epsilon}^\epsilon) h_i(g_\frac{t}{\epsilon}^\epsilon) &= f_i(\tilde{x}_\frac{r}{\epsilon}^\epsilon) h_i(g_\frac{r}{\epsilon}^\epsilon) + \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} (Df_i)_{\tilde{x}_s^\epsilon}(H_{\tilde{x}_s^\epsilon}(g_s^\epsilon e_0)) h_i(g_s^\epsilon) ds \\ &\quad + \frac{1}{\sqrt{\epsilon}} \sum_k \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} f_i(\tilde{x}_s^\epsilon) (Dh_i)_{(g_s^\epsilon)}(g_s^\epsilon A_k) dw_s^k \\ &\quad + \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} f_i(\tilde{x}_s^\epsilon) L_{g_s^\epsilon A} h_i(g_s^\epsilon) ds + \frac{1}{\epsilon} \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} f_i(\tilde{x}_s^\epsilon) \mathcal{L}_G h_i(g_s^\epsilon) ds. \end{aligned}$$

We sum up the above equation from $i = 1$ to n , since $\sum_{i=1}^n f_i(u) \mathcal{L}_G h_i(g) = H_u(g e_0)$ we identify the last term as that in $F(\tilde{x}_\frac{t}{\epsilon}^\epsilon)$. Plug this back to $F(\tilde{x}_\frac{t}{\epsilon}^\epsilon)$ to see the following.

$$\begin{aligned}
F(\tilde{x}_{\frac{t}{\epsilon}}^\epsilon) &= F(\tilde{x}_{\frac{r}{\epsilon}}^\epsilon) + \epsilon \sum_{i=1}^n \left(f_i(\tilde{x}_{\frac{t}{\epsilon}}^\epsilon) h_i(g_{\frac{t}{\epsilon}}^\epsilon) - f_i(\tilde{x}_{\frac{r}{\epsilon}}^\epsilon) h_i(g_{\frac{r}{\epsilon}}^\epsilon) \right) \\
&\quad - \epsilon \sum_{i=1}^n \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} (Df_i)_{\tilde{x}_s^\epsilon} (H_{\tilde{x}_s^\epsilon}(g_s^\epsilon e_0)) h_i(g_s^\epsilon) ds \\
&\quad - \epsilon \sum_{i=1}^n \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} f_i(\tilde{x}_s^\epsilon) L_{g_s^\epsilon \bar{A}} h_i(g_s^\epsilon) ds \\
&\quad - \sqrt{\epsilon} \sum_{i=1}^n \sum_{k=1}^N \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} f_i(\tilde{x}_s^\epsilon) (Dh_i)_{(g_s^\epsilon)} (g_s^\epsilon A_k) dw_s^k
\end{aligned}$$

To differentiate $f_i(u) = (DF)_u(H_u e_i)$ we use the flat connection ∇ on P determined by the parallelization $\mathbb{X} : P \times \mathbb{R}^n \times \mathfrak{so}(n) \rightarrow TP$ where $\mathbb{X}_u(e, A) = H_u(e) + \varpi_u^{-1}(A)$. In the calculation below we use the fact that $\nabla H(e) = 0$.

$$\begin{aligned}
&F(\tilde{x}_{\frac{t}{\epsilon}}^\epsilon) - F(\tilde{x}_{\frac{r}{\epsilon}}^\epsilon) \\
&= \epsilon \sum_{i=1}^n \left((DF)_{\tilde{x}_{\frac{t}{\epsilon}}^\epsilon} (H_{\tilde{x}_{\frac{t}{\epsilon}}^\epsilon} e_i) h_i(g_{\frac{t}{\epsilon}}^\epsilon) - (DF)_{\tilde{x}_{\frac{r}{\epsilon}}^\epsilon} (H_{\tilde{x}_{\frac{r}{\epsilon}}^\epsilon} e_i) h_i(g_{\frac{r}{\epsilon}}^\epsilon) \right) \\
(3.2) \quad &- \epsilon \sum_{i=1}^n \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} (\nabla DF)_{\tilde{x}_s^\epsilon} (H_{\tilde{x}_s^\epsilon}(g_s^\epsilon e_0), H_{\tilde{x}_s^\epsilon}(e_i)) h_i(g_s^\epsilon) ds \\
&- \epsilon \sum_{i=1}^n \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} (DF)_{\tilde{x}_s^\epsilon} (H_{\tilde{x}_s^\epsilon} e_i) L_{g_s^\epsilon \bar{A}} h_i(g_s^\epsilon) ds \\
&- \sqrt{\epsilon} \sum_{i=1}^n \sum_{k=1}^N \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} (DF)_{\tilde{x}_s^\epsilon} (H_{\tilde{x}_s^\epsilon} e_i) (Dh_i)_{(g_s^\epsilon)} (g_s^\epsilon A_k) dw_s^k.
\end{aligned}$$

We also remark that for $|H_{\tilde{x}_s^\epsilon} e_i| = 1$, $|H_{\tilde{x}_s^\epsilon} g_s^\epsilon e_i| = 1$, $|g_s^\epsilon| = |\bar{A}|$. If F is a function that is BC^2 , by the Kunita-Watanabe inequality, for any $p \geq 1$,

$$\mathbb{E} \left| F(\tilde{x}_{\frac{t}{\epsilon}}^\epsilon) - F(\tilde{x}_{\frac{r}{\epsilon}}^\epsilon) \right|^p \leq C_1(T) \epsilon^p (|DF|_\infty + |\nabla dF|_\infty) + C(T) |DF|_\infty |t-r|^{\frac{p}{2}},$$

for some constant $C_1(T)$. Hence $\mathbb{E} \left| F(\tilde{x}_{\frac{t}{\epsilon}}^\epsilon) - F(\tilde{x}_{\frac{r}{\epsilon}}^\epsilon) \right|^p \leq 2C_1(T) |t-r|^{\frac{p}{2}}$, when $\epsilon^2 \leq |t-r|$. If $|t-r| < \epsilon^2$, we estimate directly from (3.1)

$$|F(\tilde{x}_{\frac{t}{\epsilon}}^\epsilon) - F(\tilde{x}_{\frac{r}{\epsilon}}^\epsilon)| \leq C \frac{t-r}{\epsilon} \leq C \sqrt{t-r}.$$

Thus, for $C(T) = 2C_1(T) + C^p$,

$$\mathbb{E} \left| F(\tilde{x}_t^\epsilon) - F(\tilde{x}_r^\epsilon) \right|^p \leq C(T) |t - r|^{\frac{p}{2}}.$$

We apply the above formula to $F = \phi^2 \circ \tilde{\rho}(\cdot, u_0)$ where $u_0 = \tilde{x}_0^\epsilon$. Since ϕ is bounded so is F . Since $|\nabla \tilde{\rho}(\cdot, u_0)| = 1$ and ϕ' is bounded, $\nabla F = 2\phi\phi'\nabla\rho(\cdot, u_0)$ is bounded. The norm of its second derivative is:

$$|2(\phi')^2 \nabla \rho \otimes \nabla \rho + 2(\phi\phi'') \nabla \rho \otimes \nabla \rho + 2(\phi\phi') \nabla d\rho|,$$

and the tensor is evaluated at $\rho(x, y)$. We remark that $\phi'(x, y) = 0$ when $\rho(x, y) \geq a$ and $|\nabla d\rho(\rho(x, y))| \leq C$ when $\rho(x, y) \geq a$. Hence for all u_0 , there is a common number $C(T)$ s.t.

$$\mathbb{E} \left| \tilde{d}(\tilde{x}_t^\epsilon, u_0) \right|^p \leq C(T) t^{\frac{p}{2}}.$$

Conditioning on \mathcal{F}_r to see that,

$$\mathbb{E} \left| \tilde{d}(\tilde{x}_t^\epsilon, \tilde{x}_r^\epsilon) \right|^p \leq C(T) |t - r|^{\frac{p}{2}}.$$

The tightness of the law of $\{\tilde{x}_t^\epsilon\}$ follows. By Kolmogorov's criterion, $\{\tilde{x}_t^\epsilon\}$ is Hölder continuous with exponent α for any $\alpha < \frac{1}{2}$. The Hölder constants are independent of ϵ and, for any $p' < p$, Kolmogorov's criterion yields

$$(3.3) \quad \sup_{\epsilon} \mathbb{E} \sup_{s \neq t} \left(\frac{\tilde{d}(\tilde{x}_t^\epsilon, \tilde{x}_s^\epsilon)}{|t - s|^\alpha} \right)^{p'} < \infty,$$

thus concluding the proof. \square

We will need the following lemma in which we make a statement on the limit of a function of two variables, one of which is ergodic and the other one varies significantly slower. The result is straightforward, but we include the proof for completeness. If $f : N \rightarrow \mathbb{R}$ is a Lipschitz continuous function on a metric space (N, d) with distance function d , we denote by $|f|_{Lip}$ its Lipschitz semi-norm. If S is a subset of N , we let $\text{Osc}_S(f)$ denote $|\sup_{x \in S} f(x) - \inf_{x \in S} f(x)|$, the Oscillation of f over S . Let $\text{Osc}(f) = \text{Osc}_N(f)$.

Let $E(N)$ be one of the following classes of real valued functions on a metric space (N, d) ,

$$E(N) = \{f : N \rightarrow \mathbb{R} : |f|_{Lip} < \infty, \text{Osc}(f) < \infty\}$$

or $E_r(N) = E(N) \cap C^r$, where $r = 0, 1, \dots, \infty$. Denote

$$|f|_E = |f|_{Lip} + \text{Osc}(f).$$

Let d be the metric with respect to which the Lipschitz property is defined. We define $\tilde{d} = d \wedge 1$ to be a new metric on N . Then $|f|_{Lip} \leq C$ and $\text{Osc}(f) \leq C$ is equivalent to f being Lipschitz with respect to \tilde{d} .

Let $p > 1$ and let us denote the Wasserstein p -distance between two probability measures on a metric space with distance d by $W_p(N)$:

$$(W_p(\mu_1, \mu_2))^p = \inf_{\{\nu: (\pi_1)_*\nu = \mu_1, (\pi_2)_*\nu = \mu_2\}} \int_{N \times N} (d(x, y))^p d\nu(x, y).$$

Let μ^ϵ, μ be a family of probability measures on the metric space (N, d) . Then $\mu^\epsilon \rightarrow \mu$ in $W_p(N)$ if and only if they converge weakly and $\sup_{x \in N} \int (d(x, y))^p d\mu^\epsilon(y)$ is bounded for any $x \in N$. If $\tilde{d} = d \wedge 1$, then \tilde{d} and d induce the same topology on N and the concepts of weak convergence are equivalent. With respect to \tilde{d} , weak convergence is equivalent to Wasserstein p -convergence.

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ be a filtered probability space. Let $(Y, \rho), (Z, d)$ be metric spaces or C^m manifolds. Let $\{(y_t^\epsilon, t \leq T), \epsilon > 0\}$ be a family of \mathcal{F}_t -adapted stochastic processes with state space Y . Let (z_t^ϵ) be a family of sample continuous \mathcal{F}_t -Markov processes on Z .

- Assumption 3.3.** (1) The stochastic processes $(y_t^\epsilon, t \leq T)$ are equi uniformly continuous and converge weakly to a continuous process $(\bar{y}_t, t \leq T)$.
- (2) For each ϵ , $(z_t^\epsilon, t \leq T)$ has an invariant measure μ_ϵ . There exists a function C on $\mathbb{R}_+ \times Z \times \mathbb{R}_+$ with the property that $\delta(\cdot, z, \epsilon)$ is non-decreasing for each pair of (z, ϵ) and $\lim_{\epsilon \rightarrow 0} \sup_{z \in Z} \delta(K, z, \epsilon) = 0$ for all K and for all $f \in E_r(Z)$ and $t > 0$,

$$\mathbb{E} \left| \frac{\epsilon}{t} \int_0^{\frac{t}{\epsilon}} f(z_{s\epsilon}^\epsilon) ds - \int_Z f(z) d\mu_\epsilon(z) \right| \leq \delta(|f|_E, z_0^\epsilon, \frac{\epsilon}{t}).$$

- (3) There exists a probability measure μ on $W^1(C([0, T]; Z))$ s.t. $\lim_{\epsilon \rightarrow 0} W_1(\mu_\epsilon, \mu) = 0$.
- (4) The processes (y_t^ϵ) converges to (\bar{y}_t) in $W_1(Z)$, and there exists an exponent $\alpha > 0$ such that

$$\sup_{\epsilon} \mathbb{E} \left(\sup_{s \neq t} \frac{d(y_s^\epsilon, y_t^\epsilon)}{|t - s|^\alpha} \right) < \infty.$$

We cannot assume that (\bar{y}_t) is adapted to the filtration with respect to which (z_t^ϵ) is a Markov process. The process (z_t^ϵ) is usually not convergent and we do not assume that $(y_t^\epsilon, z_t^\epsilon)$ and (\bar{y}_t) are realized in the same probability space.

We denote by \hat{P}_η the probability distribution of a random variable η and let T be a positive real number. If r is a positive number, let $C([0, r]; Y)$ denotes the space of continuous paths, $\sigma : [0, r] \rightarrow Y$, on Y . If $F : C([0, r]; Y) \rightarrow \mathbb{R}$ is a Borel measurable function we use the shorter notation $F(y_\epsilon^\epsilon)$ for $F\left((y_\epsilon^\epsilon, u \leq r)\right)$.

Lemma 3.4. *Let $(\Omega, \mathcal{F}, (\mathcal{F}_t), P)$ be a filtered probability space. Let $(Y, \rho), (Z, d)$ be metric spaces or C^m manifolds in case $m \geq 1$. Let $\{(y_t^\epsilon, t \leq T), \epsilon > 0\}$ be a family of \mathcal{F}_t -adapted stochastic processes on Y . Let (z_t^ϵ) be a family of sample continuous \mathcal{F}_t -Markov processes on Z . Let $G \in E_m(Y \times Z)$. Let $0 \leq r < t$ and let $F : C([0, r]; Y) \rightarrow \mathbb{R}$ be a bounded continuous function. We define*

$$A(\epsilon) \equiv A(\epsilon, F, G) := F(y_\epsilon^\epsilon) \int_r^t G(y_\epsilon^\epsilon, z_\epsilon^\epsilon) ds.$$

- If (1)-(3) in Assumption 3.3 holds, then the random variables $A(\epsilon)$ converge weakly to A as $\epsilon \rightarrow 0$, where

$$A \equiv A(F, G) := F(\bar{y}) \int_r^t \int_Z G(\bar{y}_s, z) d\mu(z) ds.$$

- Assume (1)-(4) in Assumption 3.3. Then there is a constant c , s.t. for $\epsilon < 1$,

$$\begin{aligned} W_1\left(\hat{P}_{A(\epsilon)}, \hat{P}_A\right) &\leq c|F|_\infty \max_{z \in Z} \delta\left(|G|_E, z, \frac{\epsilon}{t-r}\right) \\ &+ c(t-r)|F|_\infty |G|_{Lip}\left(\epsilon^\alpha + W_1\left(\hat{P}_{y_\epsilon^\epsilon}, \hat{P}_{\bar{y}}\right)\right) + W_1(\mu^\epsilon, \mu). \end{aligned}$$

Proof. Let us fix the functions F, G, r, t and define:

$$\begin{aligned} \mathcal{E}_1(r, t) &= \int_r^t G(y_\epsilon^\epsilon, z_\epsilon^\epsilon) ds - \int_r^t \int_Z G(y_\epsilon^\epsilon, z) d\mu_\epsilon(z) ds; \\ \mathcal{E}_2 &= F(y_\epsilon^\epsilon) \left(\int_r^t \int_Z G(y_\epsilon^\epsilon, z) d\mu_\epsilon(z) ds - \int_r^t \int_Z G(y_\epsilon^\epsilon, z) d\mu(z) ds \right); \\ I(\epsilon) &= F(y_\epsilon^\epsilon) \int_r^t \int_Z G(y_\epsilon^\epsilon, z) d\mu(z) ds; \\ I &= F(\bar{y}) \int_r^t \int_Z G(\bar{y}_s, z) d\mu(z) ds. \end{aligned}$$

The proof is split into three parts: (i) $F(y_\epsilon^\epsilon) \mathcal{E}_1(r, t)$ converges to zero in $L_p(\Omega)$ for any $p > 1$, (ii) \mathcal{E}_2 converges to zero in $L_p(\Omega)$ for any $p > 1$, and (iii) $I(\epsilon)$ converges to I weakly.

We first prove that $F(y^\epsilon([0, \frac{r}{\epsilon}])) \int_r^t G(y_{\frac{s}{\epsilon}}^\epsilon, z_{\frac{s}{\epsilon}}^\epsilon) ds$ converges to zero in $L_p(\Omega)$. Since F is bounded and $(y_t^\epsilon, z_t^\epsilon)$ is a Markov process, it is sufficient to take $r = 0$ and F a constant, and to work with $\mathcal{E}_1(0, r)$. Let us write

$$\mathcal{E}_1 := \int_0^t G(y_{\frac{s}{\epsilon}}^\epsilon, z_{\frac{s}{\epsilon}}^\epsilon) ds - \int_0^t \int_Z G(y_{\frac{s}{\epsilon}}^\epsilon, z) d\mu_\epsilon(z) ds.$$

Let $0 = t_0 < t_1 < \dots < t_M \leq t$ be a partition of $[0, t]$ into pieces of size $t\epsilon$. Let $M \equiv M_\epsilon = [\frac{t}{t\epsilon}]$. Let $\Delta t_i = t_{i+1} - t_i$ and let $\tilde{t} = t\epsilon M_\epsilon$. Below $a \sim b$ indicates ' $a - b = o(\epsilon)$ ' as ϵ converges to 0. Since G is bounded,

$$\left| \int_{\tilde{t}}^t G(y_{\frac{s}{\epsilon}}^\epsilon, z_{\frac{s}{\epsilon}}^\epsilon) ds \right| \leq |G|_\infty(t - \tilde{t}) \sim \epsilon |G|_\infty.$$

By the Lipschitz continuity of G , for each $\epsilon > 0$ the following holds.

$$\begin{aligned} \mathcal{E}_4 &:= \left| \sum_{i=0}^{M_\epsilon-1} \int_{t_i}^{t_{i+1}} G(y_{\frac{s}{\epsilon}}^\epsilon, z_{\frac{s}{\epsilon}}^\epsilon) ds - \sum_{i=0}^{M_\epsilon-1} \int_{t_i}^{t_{i+1}} G(y_{\frac{t_i}{\epsilon}}^\epsilon, z_{\frac{s}{\epsilon}}^\epsilon) ds \right| \\ &\leq |G|_{Lip} \sum_{i=0}^{M_\epsilon-1} \int_{t_i}^{t_{i+1}} \rho(y_{\frac{s}{\epsilon}}^\epsilon, y_{\frac{t_i}{\epsilon}}^\epsilon) ds \end{aligned}$$

By equi uniform continuity of $(y_{\frac{s}{\epsilon}}^\epsilon)$, for almost surely all ω , \mathcal{E}_4 converges to zero. Since \mathcal{E}_4 is bounded the convergence is in $L_p(\Omega)$. If $(y_{\frac{s}{\epsilon}}^\epsilon)$ is assumed to be equi Hölder continuous as in condition (4), there is a convergence rate of $\epsilon^\alpha |G|_{Lip}$ for the L^p convergence.

We prove next that $\sum_{i=0}^{M_\epsilon-1} \int_{t_i}^{t_{i+1}} G(y_{\frac{t_i}{\epsilon}}^\epsilon, z_{\frac{s}{\epsilon}}^\epsilon) ds$ converges. We apply the Markov property of (z_t^ϵ) and we use the fact that (y_t^ϵ) is adapted to the filtration (\mathcal{F}_t) , with respect to which (z_t^ϵ) is a Markov process.

$$\begin{aligned} &\sum_{i=1}^{M_\epsilon-1} \mathbb{E} \left| \int_{t_i}^{t_{i+1}} G(y_{\frac{t_i}{\epsilon}}^\epsilon, z_{\frac{s}{\epsilon}}^\epsilon) ds - \Delta t_i \int_Z G(y_{\frac{t_i}{\epsilon}}^\epsilon, z) d\mu_\epsilon(z) \right| \\ &\leq \sum_{i=1}^{M_\epsilon-1} \Delta t_i \mathbb{E} \left(\mathbb{E} \left\{ \left| \frac{1}{\Delta t_i} \int_{t_i}^{t_{i+1}} G(y_{\frac{t_i}{\epsilon}}^\epsilon, z_{\frac{s}{\epsilon}}^\epsilon) ds - \int_Z G(y_{\frac{t_i}{\epsilon}}^\epsilon, z) d\mu_\epsilon(z) \right| \middle| \mathcal{F}_{\frac{t_i}{\epsilon}} \right\} \right) \\ &= \sum_{i=1}^{M_\epsilon-1} \Delta t_i \mathbb{E} \left(\mathbb{E} \left(\left| \frac{\epsilon^2}{\Delta t_i} \int_{\frac{t_i}{\epsilon^2}}^{\frac{t_{i+1}}{\epsilon^2}} G(y, z_{s\epsilon}^\epsilon) ds - \int_Z G(y, z) d\mu_\epsilon(z) \right| \middle| y = y_{\frac{t_i}{\epsilon}}^\epsilon \right) \right). \end{aligned}$$

Since $\frac{\epsilon^2}{\Delta t_i} = \frac{\epsilon}{t}$, we may now apply condition (2) and obtain

$$\begin{aligned} & \mathbb{E} \left(\left| \frac{\epsilon^2}{\Delta t_i} \int_{\frac{t_i}{\epsilon^2}}^{\frac{t_{i+1}}{\epsilon^2}} G(y, z_{s\epsilon}^\epsilon) ds - \int_Z G(y, z) d\mu_\epsilon(z) \right| \right) \\ & \leq \delta \left(\left| G(y_{\frac{t_i}{\epsilon}}^\epsilon, \cdot) \right|_E, z_{\frac{t_i}{\epsilon}}^\epsilon, \frac{\epsilon}{t} \right) \leq \delta \left(|G|_E, z_{\frac{t_i}{\epsilon}}^\epsilon, \frac{\epsilon}{t} \right). \end{aligned}$$

We record that

$$\begin{aligned} (3.4) \quad \mathcal{E}_5 &:= \mathbb{E} \left| \sum_{i=0}^{M_\epsilon-1} \int_{t_i}^{t_{i+1}} G(y_{\frac{t_i}{\epsilon}}^\epsilon, z_{\frac{s}{\epsilon}}^\epsilon) ds - \sum_{i=0}^{M_\epsilon-1} \Delta t_i \int_Z G(y_{\frac{t_i}{\epsilon}}^\epsilon, z) d\mu_\epsilon(z) \right| \\ &\leq \max_{z \in Z} \delta \left(|G|_E, z, \frac{\epsilon}{t} \right). \end{aligned}$$

Let us define

$$\mathcal{E}_6 := \sum_{i=0}^{M_\epsilon-1} \Delta t_i \int_Z G(y_{\frac{t_i}{\epsilon}}^\epsilon, z) d\mu_\epsilon(z) - \int_0^t \int_Z G(y_{\frac{s}{\epsilon}}^\epsilon, z) d\mu(z) ds.$$

By the definition of Riemann integral

$$\mathcal{E}_6 \leq |G|_{Lip} \sum_{i=0}^{M_\epsilon-1} \Delta t_i \text{Osc}_{[s_i, s_{i+1}]}(y_{\frac{s}{\epsilon}}^\epsilon)$$

where $\text{Osc}_{[a,b]}(f)$ denotes the oscillation of a function f in the indicated interval. Since $(y_{\frac{s}{\epsilon}}^\epsilon)$ is equiv uniform continuous on $[0, T]$, $\mathcal{E}_6 \rightarrow 0$ in L_p . Given Hölder continuity of $(y_{\frac{s}{\epsilon}}^\epsilon)$ from condition (4), we have the quantitative estimates: $|\mathcal{E}_6|_{L_p(\Omega)} \leq C|G|_{Lip}\epsilon^\alpha$. It follows that $F(y_{\frac{r}{\epsilon}}^\epsilon) \mathcal{E}_1(r, t) \rightarrow 0$ in L_p .

When condition (4) holds, there is a constant C such that

$$(3.5) \quad \left| F(y_{\frac{r}{\epsilon}}^\epsilon) \mathcal{E}_1(r, t) \right|_{L_p(\Omega)} \leq C(\epsilon^\alpha + \epsilon) |F|_\infty |G|_{Lip} + C|F|_\infty \max_{z \in Z} \delta \left(|G|_E, z, \frac{\epsilon}{t-r} \right).$$

For any two random variables on the same probability space and with the same state space, the L_p norm of their difference dominates their Wasserstein p -distance. The random variable

$$F(y_{\frac{r}{\epsilon}}^\epsilon) \int_r^t G(y_{\frac{s}{\epsilon}}^\epsilon, z_{\frac{s}{\epsilon}}^\epsilon) ds - F(y_{\frac{r}{\epsilon}}^\epsilon) \int_r^t \int_Z G(y_{\frac{s}{\epsilon}}^\epsilon, z) d\mu_\epsilon(z) ds \xrightarrow{W_p(N)} 0,$$

with the same rate as indicated above.

We proceed to step (ii). It is clear that for almost all ω , $F(y_\epsilon^\epsilon) \int_r^t G(y_\epsilon^\epsilon, z) ds$ is Lipschitz continuous in z . For any $z_1, z_2 \in Z$,

$$\begin{aligned} & \left| F(y_\epsilon^\epsilon) \int_r^t G(y_\epsilon^\epsilon, z_1) ds - F(y_\epsilon^\epsilon) \int_r^t G(y_\epsilon^\epsilon, z_2) ds \right| \\ & \leq |F|_\infty d(z_1, z_2) \int_r^t |G(y_\epsilon^\epsilon, \cdot)|_{Lip} ds \leq (t-r) d(z_1, z_2) |F|_\infty |G|_{Lip}. \end{aligned}$$

By the Kantorovich duality formula, for the distance between two probability measures μ_1 and μ_2 :

$$W_1(\mu_1, \mu_2) = \sup \left\{ \int U d\mu_1 - \int U d\mu_2 : |U|_{Lip} \leq 1 \right\},$$

we have

$$|\mathcal{E}_2| \leq (t-r) \cdot |F|_\infty \cdot |G|_{Lip} \cdot W_1(\mu^\epsilon, \mu).$$

For part (iii) let U be a continuous function on $C([0, T]; Y)$. If $\sigma \in C([0, T]; Y)$, let us denote by $\sigma([0, r])$ the restriction of the path to $[0, r]$. Since F is bounded continuous and G is Lipschitz continuous,

$$\sigma \mapsto U \left(F(\sigma([0, r])) \left(\int_r^t \int_Z G(\sigma_s, z) d\mu(z) ds \right) \right)$$

is a continuous function on $C([0, T]; Y)$. By the weak convergence of (y_ϵ^ϵ) , $\mathbb{E}(U(I(\epsilon)))$ converges to $\mathbb{E}(U(I))$ and the random variables $I(\epsilon)$ converge weakly to I . By now we have proved that $A(\epsilon, F, G)$ converges to $A(F, G)$ weakly; we thus conclude the first part of the lemma.

Let us assume condition (4) from Assumption 3.3. In particular (y_ϵ^ϵ) converges in $W_1(C([0, T]; Y))$. Let U be a Lipschitz continuous function on $C([0, T]; Y)$. We define $\tilde{U} : C([0, T]; Y) \rightarrow \mathbb{R}$ by

$$\tilde{U}(\sigma) = U \left(F(\sigma([0, r])) \left(\int_r^t \int_Z G(\sigma_s, z) d\mu(z) ds \right) \right).$$

Let σ^1, σ^2 are two paths on Y ,

$$\begin{aligned} & \left| \tilde{U}(\sigma_1) - \tilde{U}(\sigma_2) \right| \\ & \leq |U|_{Lip} \cdot |F|_\infty \left| \int_r^t \int_Z G(\sigma_s^1, z) d\mu(z) ds - \int_r^t \int_Z G(\sigma_s^2, z) d\mu(z) ds \right| \\ & \leq (t-r) |U|_{Lip} \cdot |F|_\infty \cdot |G|_{Lip} \cdot \sup_{0 \leq s \leq T} \rho(\sigma_s^1, \sigma_s^2). \end{aligned}$$

By Kantorovich duality and Assumption (4),

$$W_1 \left(\hat{P}_{I(\epsilon)}, \hat{P}_I \right) \leq (t-r) \cdot |F|_\infty \cdot |G|_{Lip} \cdot W_1 \left(\hat{P}_{y_\epsilon^\epsilon}, \hat{P}_{\bar{y}} \right).$$

We collect all the estimations together. Under Assumption (1-4), the following estimates hold.

$$\begin{aligned} W_1\left(\hat{P}_{A(\epsilon)}, \hat{P}_A\right) &\leq C|F|_\infty |G|_{Lip}(\epsilon^\alpha + \epsilon) + C|F|_\infty \max_{z \in Z} \delta\left(|G|_E, z, \frac{\epsilon}{t-r}\right) \\ &+ C(t-r) \cdot |F|_\infty \cdot |G|_{Lip} \cdot \left(W_1\left(\hat{P}_{y_\epsilon}, \hat{P}_{\bar{y}}\right) + W_1(\mu_\epsilon, \mu)\right). \end{aligned}$$

We may now limit ourselves to $\epsilon \leq 1$ and conclude part (2) of the Lemma. \square

Remark 3.5. In the lemma above we should really think that the z^ϵ process and process y^ϵ follow different clocks, the former is run at the fast time scale $\frac{1}{\epsilon}$ and the latter at scale 1.

Example 3.6. Let (g_s) be a Brownian motion on $G = SO(n)$, solving

$$dg_t = \sum_{k=1}^N L_{g_t A_k} dw_t^k.$$

Here $\{A_1, \dots, A_N\}$ is an orthonormal basis of \mathfrak{g} . In Lemma 3.4 we take $z_t^\epsilon = g_{\frac{t}{\epsilon}}$, then condition (2) holds. If f is a Lipschitz continuous function, it is well known that the law of large numbers holds for $\int_0^t f(g_s) ds$, so does a central limit theorem. The remainder term in the central limit theorem is of order \sqrt{t} and depends on f only through the Lipschitz constant $|f|_{Lip}$.

It is easy to see that the remainder term in the law of large numbers depends only on the Lipschitz constant of the function. Without loss of generality we assume that $\int f dg = 0$. Let α solve the Poisson equation: $\Delta^G \alpha = f$. Then

$$\frac{1}{t} \int_0^t f(g_s) ds = \frac{1}{t} \alpha(g_t) - \frac{1}{t} \alpha(g_0) - \sum_k \frac{1}{t} \int_0^t (D\alpha)(g_s A_k) dw_s^k.$$

Since α is bounded, we are only concerned with the martingale term. By Burkholder-Davis-Gundy inequality, its L^2 norm is bounded by

$$\frac{1}{t} \left(\sum_{k=1}^N \int_0^t \mathbb{E}((D\alpha)(g_s A_k))^2 \right)^{\frac{1}{2}} ds \leq \frac{\sqrt{N}}{\sqrt{t}} \int_0^t \mathbb{E}|D\alpha|_{g_s}^2 ds.$$

By elliptic estimates, $|D\alpha|$ is bounded by the $|f|_{L_\infty}$. Since f is centred, it is bounded by $\text{Osc}(f)$. In summary,

$$\mathbb{E} \left(\frac{1}{t} \int_0^t f(g_s) ds - \int_N f(g) dg \right)^2 \leq \sqrt{N} \text{Osc}(f) t^{-\frac{1}{2}}.$$

In Theorem 1.1 we may wish to add an extra drift of the form $\frac{1}{\epsilon}A^*$ where $A \in \mathfrak{g}$, so that \mathcal{L}_G is $\frac{1}{2}\Delta^G + L_{gA}$. Translations by orthogonal matrices are isometries, so for any $A \in \mathfrak{g}$ the vector field gA is a killing field, and the Haar measure remains an invariant measure for the diffusion with infinitesimal generator $\frac{1}{2}\Delta^G + L_{gA}$. However, on a compact Lie group no left invariant vector field is the gradient of a function and $\frac{1}{2}\Delta^G + L_{gA}$ is no longer a symmetric operator. In this case we do not know how to obtain the estimate in the example.

4. PROOF

We are ready to prove the main theorem. We use some ideas from Papanicolaou, Stroock, Varadhan [38, 37] and Hairer, Pavliotis [19].

Proof. We define a Markov generator \bar{L} on OM . If $F : OM \rightarrow \mathbb{R}$ is bounded and Borel measurable and $\{e_i\}$ is an orthonormal basis of \mathbb{R}^n , we define

$$(4.1) \quad \begin{aligned} \bar{L}F = & - \sum_{i=1}^n \int_G (\nabla DF)_u(H_u(ge_0), H_u(e_i)) h_i(g) dg \\ & - \sum_{i=1}^n \int_G (DF)_u(H_u e_i) L_{gA} h_i(g) dg. \end{aligned}$$

Since (\tilde{x}_t^ϵ) is tight by Lemma 3.2, every sub-sequence of (\tilde{x}_t^ϵ) has a sub-sequence that converges in distribution. We will prove that the probability distributions of (\tilde{x}_t^ϵ) converge weakly to the probability measure, \bar{P} , determined by \bar{L} . It is sufficient to prove that if (\bar{y}_t) is a limit of (\tilde{x}_t^ϵ) , then

$$F(\bar{y}_t) - F(u_0) - \int_0^t \bar{L}F(\bar{y}_s) ds$$

is a martingale. Since the convergence is weak, and the Markov process $(\tilde{x}_t^\epsilon, g_t^\epsilon)$ is not tight, we do not have a suitable filtration on Ω to work with. We formulate the above convergence on the space of continuous paths over OM on a given time interval $[0, T]$.

Let X_t be the coordinate process on the path space over OM , $\mathcal{G}_t = \sigma\{(X_s) : 0 \leq s \leq t\}$ and let $\hat{P}_{\tilde{x}^\epsilon}$ be the probability distribution of (\tilde{x}_t^ϵ) on the path space over OM . By taking a subsequence if necessary, we may assume that $\{\hat{P}_{\tilde{x}^\epsilon}\}$ converges to \bar{P} .

Let $F : OM \rightarrow \mathbb{R}$ be a smooth function with compact support. We will prove that with respect to \bar{P} ,

$$\mathbb{E} \left\{ (F(X_t) - F(X_r) - \int_r^t \bar{\mathcal{L}}F(X_s)ds \mid \mathcal{G}_r) \right\} = 0.$$

Since $\hat{P}_{\tilde{x}_\epsilon} \rightarrow \bar{P}$ weakly, we only need to prove that for all bounded and continuous real value random variables ξ that are measurable with respect to \mathcal{G}_r ,

$$(4.2) \quad \lim_{\epsilon \rightarrow 0} \int \xi(F(X_t) - F(X_r))d\hat{P}_{\tilde{x}^\epsilon} = \mathbb{E} \left(\xi \int_r^t \bar{L}F(X_s)dsd\bar{P} \right).$$

By formula (3.2) in the proof of Lemma 3.2, for $t \geq r$,

$$(4.3) \quad \begin{aligned} & F(\tilde{x}_{\frac{t}{\epsilon}}^\epsilon) - F(\tilde{x}_{\frac{r}{\epsilon}}^\epsilon) \\ & \sim -\epsilon \sum_{i=1}^n \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} (\nabla DF)_{\tilde{x}_s^\epsilon} (H_{\tilde{x}_s^\epsilon}(g_s^\epsilon e_0), H_{\tilde{x}_s^\epsilon}(e_i)) h_i(g_s^\epsilon) ds \\ & - \epsilon \sum_{i=1}^n \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} (DF)_{\tilde{x}_s^\epsilon} (H_{\tilde{x}_s^\epsilon} e_i) L_{g_s^\epsilon A} h_i(g_s^\epsilon) ds \\ & - \sqrt{\epsilon} \sum_{i=1}^n \sum_{k=1}^N \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} (DF)_{\tilde{x}_s^\epsilon} (H_{\tilde{x}_s^\epsilon} e_i) (Dh_i)_{(g_s^\epsilon)} (g_s^\epsilon A_k) dw_s^k. \end{aligned}$$

Hence up to a term of order ϵ ,

$$\begin{aligned} & \int \xi(F(X_t) - F(X_r)) d\hat{P}_{\tilde{x}^\epsilon} \\ & = o(\epsilon) - \epsilon \sum_{i=1}^n \int \left(\xi \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} (\nabla DF)_{X_s} (H_{X_s}(G_s e_0), H_{X_s}(e_i)) h_i(G_s) ds \right) d\hat{P}_{\tilde{x}^\epsilon} \\ & - \epsilon \sum_{i=1}^n \int \left(\xi \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} (DF)_{X_s} (H_{X_s} e_i) L_{G_s A} h_i(G_s) ds \right) d\hat{P}_{\tilde{x}^\epsilon}. \end{aligned}$$

We prove this by working with the original processes. Let (\tilde{x}_l^ϵ) denote a sub-sequence of the original sequence with limit (\bar{y}_s) . For each $i, l = 1, \dots, n$, let us define

$$\beta_{li}(u) = (\nabla DF)_u (H_u(e_l), H_u(e_i)).$$

By linearity of H_u and ∇DF ,

$$\begin{aligned} & (\nabla DF)_u (H_u(ge_0), H_ue_i) h_i(g) \\ &= \sum_{l=1}^n (\nabla DF)_u (H_u(e_l), H_u(e_i)) \langle ge_0, e_l \rangle h_i(g) \\ &= \sum_{l=1}^n \beta_{li}(u) \langle ge_0, e_l \rangle h_i(g), \end{aligned}$$

for each $i = 1, \dots, n$; and

$$\begin{aligned} & -\epsilon \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} (\nabla DF)_{\tilde{x}_s^\epsilon} (H_{\tilde{x}_s^\epsilon}(g_s^\epsilon e_0), H_{\tilde{x}_s^\epsilon}(e_i)) h_i(\tilde{x}_s^\epsilon) ds \\ &= -\epsilon \sum_{l=1}^n \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} \beta_{li}(\tilde{x}_s^\epsilon) \langle g_s^\epsilon e_0, e_l \rangle h_i(g_s^\epsilon) ds \\ &= -\sum_{l=1}^n \int_r^t \beta_{li}(\tilde{x}_{\frac{s}{\epsilon}}^\epsilon) \langle g_{\frac{s}{\epsilon}}^\epsilon e_0, e_l \rangle h_i(g_{\frac{s}{\epsilon}}^\epsilon) ds. \end{aligned}$$

We observe that $(g_{s\epsilon}^\epsilon)$ satisfies the equation $dg_t = \sum_k g_t A_k \circ dw_t^k$ with initial value the identity. The solution stays in the connected component $SO(n)$. It is ergodic with the normalized Haar measure dg on $SO(n)$ its invariant measure and satisfies Birkhoff Ergodic Theorem, see Example 3.6. By Lemma 3.2, $(\tilde{x}_{\frac{s}{\epsilon}}^\epsilon)$ is tight, and equi uniformly Hölder continuous on $[0, T]$. In Assumption 3.3, we take $z_t^\epsilon = g_t^\epsilon$, $d\mu_\epsilon = dg$, $y_t^\epsilon = \tilde{x}_t^\epsilon$ and check that conditions (1)-(4) are satisfied. In Lemma 3.4, we take $G(x, g) = \sum_{l=1}^n \beta_{li}(u) \langle ge_0, e_l \rangle h_i(g)$. Since the function $h_i : G \rightarrow \mathbb{R}$ are smooth and G is compact, also β_{li} are smooth and bounded by construction, we may apply Lemma 3.4. If ϕ is a bounded real valued continuous function on $C([0, r]; OM)$, let $\xi = \phi(\tilde{x}_{\frac{u}{\epsilon}}^\epsilon, 0 \leq u \leq r)$. Then

$$\begin{aligned} & \lim_{\epsilon \rightarrow 0} \mathbb{E} \left(\xi \sum_{l=1}^n \int_r^t \beta_{li}(\tilde{x}_{\frac{s}{\epsilon}}^\epsilon) \langle g_{\frac{s}{\epsilon}}^\epsilon e_0, e_l \rangle h_i(g_{\frac{s}{\epsilon}}^\epsilon) ds \right) \\ &= \sum_{l=1}^n \mathbb{E} \left(\xi \int_r^t \beta_{li}(\bar{y}_s) ds \right) \int_G \langle ge_0, e_l \rangle h_i(g) dg \\ &= \sum_{l=1}^n \mathbb{E} \left(\xi \int_r^t \nabla DF_{\bar{y}_s} (H_{\bar{y}_s}(e_l), H_{\bar{y}_s}(e_i)) \right) \int_G \langle ge_0, e_l \rangle h_i(g) dg \\ &= \sum_{l=1}^n \mathbb{E} \left(\xi \int_r^t \int_G \nabla DF_{\bar{y}_s} (H_{\bar{y}_s}(ge_0), H_{\bar{y}_s}(e_i)) h_i(g) dg \right). \end{aligned}$$

For the same reasoning, we also have,

$$\begin{aligned} & \lim_{\epsilon \rightarrow 0} \epsilon \mathbb{E} \left(\xi \int_{\frac{r}{\epsilon}}^{\frac{t}{\epsilon}} (DF)_{\tilde{x}_s^\epsilon} \left(H_{\tilde{x}_s^\epsilon} e_i \right) L_{g_s^\epsilon \bar{A}} h_i(g_s^\epsilon) ds \right) \\ &= \mathbb{E} \left(\xi \int_r^t (DF)_{\bar{y}_s} (H_{\bar{y}_s} e_i) ds \int_G L_{g\bar{A}} h_i(g) dg \right). \end{aligned}$$

We have proved (4.2). Since every sub-sequence of $\hat{P}_{\tilde{x}^\epsilon}$ has a sub-sequence that converges to the same limit, we have proved $\hat{P}_{\tilde{x}^\epsilon} \rightarrow \bar{P}$ weakly. To identify the limit $\bar{\mu}$ we take $G = SO(n)$. For $g \in G$, we define

$$h_i(g) = -\frac{4}{n-1} \langle g e_0, e_i \rangle.$$

Let us first work on the first order term,

$$L_{g\bar{A}} h_i = -\frac{4}{n-1} \langle g \bar{A} e_0, e_i \rangle$$

and compute the following integral:

$$\begin{aligned} & \int_G (DF)_u (H_u e_i) L_{g\bar{A}} h_i(g) dg \\ &= -\frac{4}{n-1} \int_G (DF)_u (H_u g \bar{A} e_0) dg \\ &= -\frac{4}{n-1} (DF)_u \left(H_u \left(\int_G g \bar{A} e_0 dg \right) \right) = 0. \end{aligned}$$

Next we compute

$$\mathcal{L}_G h_i = \frac{1}{2} \sum_{k=1}^N L_{gA_k} L_{gA_k} h_i = -\frac{2}{n-1} \sum_{k=1}^N \langle g(A_k)^2 e_0, e_i \rangle.$$

Since $\sum_{k=1}^N (A_k)^2 = -\frac{n-1}{2} I$, we conclude that $\mathcal{L}_G h_i = \langle g e_0, e_i \rangle$. Let us define

$$\begin{aligned} a_{i,j}(e_0) &= - \int_G \langle g e_0, e_j \rangle \frac{1}{2} (\Delta^G)^{-1} \langle g e_0, e_i \rangle dg \\ &= \frac{4}{n-2} \int_G \langle g e_0, e_j \rangle \langle g e_0, e_i \rangle dg. \end{aligned}$$

We first prove that $a_{i,j}(e_0)$ is independent of e_0 . Let $e'_0 \in \mathbb{R}^n$ we take O such that $O e'_0 = e_0$. By the right invariant property of the Haar measure,

$$\int_G \langle g e'_0, e_j \rangle \langle g e'_0, e_i \rangle dg = \int_G \langle g O e'_0, e_j \rangle \langle g O e'_0, e_i \rangle dg = \int_G \langle g e_0, e_j \rangle \langle g e_0, e_i \rangle dg.$$

We first compute the case of $i \neq j$ and $n = 2$.

$$a_{1,2}(e_1) = \int_{SO(2)} \langle ge_1, e_2 \rangle \langle ge_1, e_2 \rangle dg = - \int_0^{2\pi} \cos(\theta) \sin(\theta) d\theta = 0.$$

If $n > 2$, for any $i \neq j$, there is an element $O \in G$ such that $Oe_i = -e_i$ and $Oe_j = e_j$. For example if $i = 1, j = 2$, we take $O = (-e_1, e_2, -e_3, e_4, \dots, e_n)$. So

$$\begin{aligned} \int_G \langle ge_0, e_j \rangle \langle ge_0, e_i \rangle dg &= - \int_G \langle ge_0, Oe_j \rangle \langle ge_0, Oe_i \rangle dg \\ &= - \int_G \langle ge_0, e_j \rangle \langle ge_0, e_i \rangle dg. \end{aligned}$$

Thus $a_{i,j} = 0$ if $i \neq j$. Let

$$C_i = \int_G \langle ge_0, e_i \rangle^2 dg.$$

For $i = 1, \dots, n$, $C_i = \int_G \langle ge_0, e_i \rangle^2 dg$ is independent of i and

$$\int_G \sum_{i=1}^n \langle ge_0, e_i \rangle^2 dg = 1$$

and consequently $C_i = \frac{1}{n}$. The non-zero values of $(a_{i,j})$ are:

$$a_{i,i} = \frac{4}{n-1} \int_G \langle ge_0, e_i \rangle^2 dg = \frac{4}{(n-1)n}.$$

By the definition, $\Delta_H F = \sum_{i=1}^n L_{H(e_i)} L_{H(e_i)} F$; we see that

$$\begin{aligned} \bar{L}F &= - \sum_{i,j} \int_G \nabla dF(H(u)e_j, H(u)e_i) \langle ge_0, e_j \rangle \mathcal{L}_G^{-1} \langle ge_0, e_i \rangle dg \\ &= \frac{4}{(n-1)n} \sum_{i=1}^n \nabla dF(\mathfrak{h}_u(ue_i), \mathfrak{h}_u(ue_i)) \\ &= \frac{4}{(n-1)n} \Delta_H. \end{aligned}$$

We conclude that $(\tilde{x}_{\frac{t}{\epsilon}}^\epsilon)$ is a diffusion process with infinitesimal generator $\frac{4}{(n-1)n} \Delta_H$. Since $(x_{\frac{t}{\epsilon}}^\epsilon)$ is the projection of $(\tilde{x}_{\frac{t}{\epsilon}}^\epsilon)$ and it is also convergent. The operators Δ^H and Δ are intertwined by π ; for $f : M \rightarrow \mathbb{R}$ smooth, $(\Delta_H f) \circ \pi = \Delta(f \circ \pi)$. See e.g. Theorem 4C of Chapter II in Elworthy [10] and also Elworthy, LeJan, Li [13]; Δ_H is cohesive and a horizontal operator in the terminology of [13] and is the horizontal lift of Δ . We see that $(x_{\frac{t}{\epsilon}}^\epsilon)$ converges to a process with generator $\frac{4}{(n-1)n} \Delta$. We have completed the proof of Theorem 1.1. \square

Acknowledgement. It is a pleasure to thank M. Hairer and J. Norris for helpful discussions; K. D. Elworthy for helpful conversations and inspiring reference. I also like to acknowledge a helpful conversation with D. Bakry and M. Ledoux.

REFERENCES

- [1] Robert Azencott. Behavior of diffusion semi-groups at infinity. *Bull. Soc. Math. France*, 102:193–240, 1974.
- [2] A. Bensoussan, J.-L. Lions, and G. Papanicolaou. *Asymptotic analysis for periodic structures*. AMS Chelsea Publishing, Providence, RI, 2011.
- [3] Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons Inc., New York, 1968.
- [4] Jean-Michel Bismut. *Hypoelliptic Laplacian and orbital integrals*, volume 177 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 2011.
- [5] A. N. Borodin and M. I. Freidlin. Fast oscillating random perturbations of dynamical systems with conservation laws. *Ann. Inst. H. Poincaré Probab. Statist.*, 31(3):485–525, 1995.
- [6] J. M. C. Clark. An introduction to stochastic differential equations on manifolds. In *Geometric Methods in System Theory, NATO Advanced Study Institutes Series*, volume 3, pages 131–149. 1973.
- [7] R. M. Dowell. *Differentiable Approximations to Brownian Motion on Manifolds*. PhD thesis, University of Warwick, 1980.
- [8] E. B. Dynkin. Diffusion of tensors. *Dokl. Akad. Nauk SSSR*, 179:1264–1267, 1968.
- [9] J. Eells and K. D. Elworthy. Stochastic dynamical systems. In *Control theory and topics in functional analysis (Internat. Sem., Internat. Centre Theoret. Phys., Trieste, 1974)*, Vol. III, pages 179–185. Internat. Atomic Energy Agency, Vienna, 1976.
- [10] David Elworthy. Geometric aspects of diffusions on manifolds. In *École d’Été de Probabilités de Saint-Flour XV–XVII, 1985–87*, volume 1362 of *Lecture Notes in Math.*, pages 277–425. Springer, Berlin, 1988.
- [11] K. D. Elworthy. Stochastic dynamical systems and their flows. In *Stochastic analysis (Proc. Internat. Conf., Northwestern Univ., Evanston, Ill., 1978)*, pages 79–95. Academic Press, New York, 1978.
- [12] K. D. Elworthy. *Stochastic differential equations on manifolds*, volume 70 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 1982.
- [13] K. David Elworthy, Yves Le Jan, and Xue-Mei Li. *The geometry of filtering*. Frontiers in Mathematics. Birkhäuser Verlag, Basel, 2010.
- [14] Michel Émery. *Stochastic calculus in manifolds*. Universitext. Springer-Verlag, Berlin, 1989. With an appendix by P.-A. Meyer.
- [15] N. Enriquez, J. Franchi, and Y. Le Jan. Central limit theorem for the geodesic flow associated with a Kleinian group, case $\delta > d/2$. *J. Math. Pures Appl. (9)*, 80(2):153–175, 2001.

- [16] Stewart N. Ethier and Thomas G. Kurtz. *Markov processes*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986. Characterization and convergence.
- [17] M. I. Freidlin and A. D. Wentzell. *Random perturbations of dynamical systems*, volume 260. Springer-Verlag, New York, second edition, 1998. Translated from the 1979 Russian original by Joseph Szücs.
- [18] Y. Guivarc'h and Y. Le Jan. Asymptotic winding of the geodesic flow on modular surfaces and continued fractions. *Ann. Sci. École Norm. Sup. (4)*, 26(1):23–50, 1993.
- [19] M. Hairer and G. A. Pavliotis. Periodic homogenization for hypoelliptic diffusions. *J. Statist. Phys.*, 117(1-2):261–279, 2004.
- [20] R. Z. Has'minskii. On the principle of averaging the Itô's stochastic differential equations. *Kybernetika (Prague)*, 4:260–279, 1968.
- [21] Inge S. Helland. Central limit theorems for martingales with discrete or continuous time. *Scand. J. Statist.*, 9(2):79–94, 1982.
- [22] Nobuyuki Ikeda. Limit theorems for a class of random currents. In *Probabilistic methods in mathematical physics (Katata/Kyoto, 1985)*, pages 181–193. Academic Press, Boston, MA, 1987.
- [23] Nobuyuki Ikeda and Yoko Ochi. Central limit theorems and random currents. In *Stochastic differential systems (Bad Honnef, 1985)*, volume 78 of *Lecture Notes in Control and Inform. Sci.*, pages 195–205. Springer, Berlin, 1986.
- [24] Kiyoshi Itô. The Brownian motion and tensor fields on Riemannian manifold. In *Proc. Internat. Congr. Mathematicians (Stockholm, 1962)*, pages 536–539. Inst. Mittag-Leffler, Djursholm, 1963.
- [25] Kiyosi Itô. Stochastic differential equations in a differentiable manifold. *Nagoya Math. J.*, 1:35–47, 1950.
- [26] Kiyosi Itô. Stochastic parallel displacement. In *Probabilistic methods in differential equations (Proc. Conf., Univ. Victoria, Victoria, B.C., 1974)*, pages 1–7. Lecture Notes in Math., Vol. 451. Springer, Berlin, 1975.
- [27] Paulo R. Ruffino Ivan I. Gonzales Gargate. An averaging principle for diffusions in foliated spaces. <http://arxiv.org/abs/1212.1587>, 2012.
- [28] R. Z. Khas'minskii. The behavior of a self-oscillating system acted upon by slight noise. *J. Appl. Math. Mech.*, 27:1035–1044, 1963.
- [29] C. Kipnis and S. R. S. Varadhan. Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions. *Comm. Math. Phys.*, 104(1):1–19, 1986.
- [30] Shoshichi Kobayashi and Katsumi Nomizu. *Foundations of differential geometry. Vol I*. Interscience Publishers, a division of John Wiley & Sons, New York-London, 1963.
- [31] Xue-Mei Li. Properties at infinity of diffusion semigroups and stochastic flows via weak uniform covers. *Potential Anal.*, 3(4):339–357, 1994.
- [32] Xue-Mei Li. An averaging principle for a completely integrable stochastic Hamiltonian system. *Nonlinearity*, 21(4):803–822, 2008.
- [33] Xue-Mei Li. Effective diffusions with intertwined structures. Preprint, 2012.
- [34] Paul Malliavin. *Géométrie différentielle stochastique*, volume 64 of *Séminaire de Mathématiques Supérieures [Seminar on Higher Mathematics]*. Presses de

- l'Université de Montréal, Montreal, Que., 1978. Notes prepared by Danièle Dehen and Dominique Michel.
- [35] Shojiro Manabe and Yoko Ochi. The central limit theorem for current-valued processes induced by geodesic flows. *Osaka J. Math.*, 26(1):191–205, 1989.
 - [36] Edward Nelson. *Dynamical theories of Brownian motion*. Princeton University Press, Princeton, N.J., 1967.
 - [37] G. C. Papanicolaou, D. Stroock, and S. R. S. Varadhan. Martingale approach to some limit theorems. In *Papers from the Duke Turbulence Conference (Duke Univ., 1976)*, pages ii+120 pp. Duke Univ., Durham, N.C., 1977.
 - [38] G. C. Papanicolaou and S. R. S. Varadhan. A limit theorem with strong mixing in Banach space and two applications to stochastic differential equations. *Comm. Pure Appl. Math.*, 26:497–524, 1973.
 - [39] Mark A. Pinsky. Homogenization and stochastic parallel displacement. In *Stochastic integrals (Proc. Sympos., Univ. Durham, Durham, 1980)*, volume 851 of *Lecture Notes in Math.*, pages 271–284. Springer, Berlin, 1981.
 - [40] M. Ratner. The central limit theorem for geodesic flows on n -dimensional manifolds of negative curvature. *Israel J. Math.*, 16:181–197, 1973.
 - [41] Ja. G. Sinai. The central limit theorem for geodesic flows on manifolds of constant negative curvature. *Soviet Math. Dokl.*, 1:983–987, 1960.

MATHEMATICS INSTITUTE, THE UNIVERSITY OF WARWICK, COVENTRY CV4 7AL, U.K., EMAIL ADDRESS: XUE-MEI.LI@WARWICK.AC.UK